



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 135120

TO: Nita M Minnifield
Location: REM-3C01-3C18
Art Unit: 1645
Tuesday, October 19, 2004

Case Serial Number: 10/034623

From: Edward Hart
Location: Biotech-Chem Library
REM-1A55
Phone: 571-272-2512

edward.hart@uspto.gov

Search Notes

Examiner Minnifield,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

*Reviewed
12/15/04*

135/20

STIC-Biotech/ChemLib

From: Minnifield, Nita
Sent: Friday, October 15, 2004 8:41 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

10/034623

STIC

Please do a commercial and interference sequence search on SEQ ID NO: 6 of the above application.

Please provide a paper copy of the results.

Thanks,
Minnifield
Art Unit 1645
Office REM-3C01
Mailbox REM-3C18
571-272-0860

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 10/18/04
Date Completed: 10/19/04
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # 1
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Q2P
WWW/Internet: _____
Other(Specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 14:31:31 ; Search time 189 Seconds

(without alignments)
818,919 Million cell updates/sec

Title: US-10-034-623-6

Sequence: 1 MHGIEGRDMSNFVAFVCACARGTGKEMKTVDRVHKECHAHGQIRPEPNEVEQ
.....RPAKKTGRRARRAGAKRR 269

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_sprot;*
2: uniprot_crembl;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1357	100.0	269	Q74057	Q74057 cenarchaeum
2	220.5	16.2	221	HC2X_CHLTR	Q46397 chlamydia t
3	219	16.1	207	HCT2_CHLTR	P38020 chlamydia m
4	215	15.8	225	Q7WFA2	Q7WFA2 bordetella
5	208	15.3	376	Q82KM1	Q82KM1 streptomyce
6	204	15.0	201	HC2D_CHLTR	Q66280 chlamydia t
7	203.5	15.0	172	HCT2_CHLTR	Q32819 chlamydia p
8	201.5	14.8	197	Q7W3X2	Q7W3X2 bordetella
9	201	14.8	328	SSS3_DROBR	P13730 diosiphilla
10	200.5	14.8	175	Q7VUT9	Q7VUT9 bordetella
11	199.5	14.7	182	Q45370	Q45370 bordetella
12	198	14.6	177	Q6SG84	Q6SG84 uncultured
13	198	14.6	177	AA37978	AA37978 uncultured
14	196.5	14.5	454	Q6IT77	Q6IT77 spleula sol
15	194.5	14.3	155	Q8P6U0	Q8P6U0 xanthomonas
16	193	14.2	156	Q8P6U0	Q8P6U0 xanthomonas
17	192	14.1	455	Q6IT76	Q6IT76 spleula sol
18	190.5	14.0	200	Q8XVY7	Q8XVY7 ralsomonia s
19	187.5	13.8	152	Q82J38	Q82J38 chlamydia p
20	186	13.7	152	Q9XYV7	Q9XYV7 euplotes cr
21	177	13.0	217	Q8ZUS0	Q8ZUS0 streptomyce
22	176.5	13.0	273	Q8PFR2	Q8PFR2 xanthomonas
23	174.5	12.9	273	Q7VAY8	Q7VAY8 prochloroc
24	174	12.8	422	HC2B_CHLTR	Q06281 chlamydia t
25	172	12.7	422	Q86G81	Q86G81 anophelies s
26	172	12.7	753	Q6PFP4	Q6PFP4 brachydanio
27	169	12.5	218	AAH57473	AAH57473 brachydan
28	169	12.5	218	DBH2_STRCO	DBH2_STRCO streptomyce
29	169	12.5	966	Q74718	Q74718 geobacter s
30	169	12.5	966	AA36668	AA36668 geobacter
31	163.5	12.0	505	Q24898	Q24898 ensis minor

32	163	12.0	371	2	Q6N4V4	Q6N4V4 rhodopseudo
33	163	12.0	371	2	CAE28670	CAE28670 rhodopseu
34	160.5	11.8	356	2	Q82AC8	Q82AC8 streptomyce
35	160.5	11.8	435	2	Q7RTL5	Q7RTL5 anophelies g
36	159.5	11.8	329	2	Q8BRD8	Q8BRD8 pseudomonas
37	158	11.6	237	2	Q91S26	Q91S26 regina rana
38	158	11.6	275	2	Q6YH51	Q6YH51 ambystoma t
39	158	11.6	275	2	MAP33252	MAP33252 ambystoma
40	156.5	11.5	259	2	Q8P419	Q8P419 xanthomonas
41	156	11.5	182	2	Q8RL36	Q8RL36 burkholderi
42	155.5	11.5	379	2	Q7WFN5	Q7WFN5 bordetella
43	155.5	11.5	212	2	Q73V19	Q73V19 mycobacteri
44	155.5	11.5	212	2	AA50572	AA50572 mycobacte
45	155	11.4	209	2	Q8PBM2	Q8PBM2 xanthomonas

ALIGNMENTS

RESULT 1

ID	Q74057	PRELIMINARY;	PRT;	269 AA.
AC	Q74057;			
DT	01-NOV-1998 (TRENBLREL.08, Created)			
DT	01-NOV-1998 (TRENBLREL.08, Last sequence update)			
DT	01-MAR-2004 (TRENBLREL.26, Last annotation update)			
DE	Histone H1 DNA binding protein.			
GN	Name=hc2;			
OS	Cenarchaeum symbiosum.			
OC	Archaea; Crenarchaeota; Thermoprotei; Cenarchaeales; Cenarchaeaceae;			
CC	Cenarchaeum.			
OX	NCBI_TaxID=46770;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STEAIN=B;			
RX	MEDLINE=98422450; PubMed=9748430;			
RA	Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,			
RA	Swanson R.V.;			
RT	"Genomic analysis reveals chromosomal variation in natural populations			
RT	of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";			
RL	J. Bacteriol. 180:5003-5009 (1998).			
DR	EMBL; AF083072; AAC62700.1; -			
DR	PIR; T31309; T31309.			
DR	SEQUENCE 269 AA; 29889 MW; ADCADCT50DAGCB8 CRC64;			

Query Match 100.0%; Score 1357; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 5.1e-91;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MHGIEGRDMSNFVAFVCACARGTGKEMKTVDRVHKECHAHGQIRPEPNEVEQ	60
DB	1	MHGIEGRDMSNFVAFVCACARGTGKEMKTVDRVHKECHAHGQIRPEPNEVEQ	60
QY	61	RVAELEKVDLIQNRNOIAEMNRASGDGVSATSAEAEOHAELEKVOIYQNRNOIAEMN	120
DB	61	RVAELEKVDLIQNRNOIAEMNRASGDGVSATSAEAEOHAELEKVOIYQNRNOIAEMN	120
QY	121	RAAPGKPAKKAAGTARRKSGKTVRRKTKGAKGAKARRKTTAKKA	180
DB	121	RAAPGKPAKKAAGTARRKSGKTVRRKTKGAKGAKARRKTTAKKA	180
QY	181	AGKAGARRKATVTKVHKIGVRRKTTARRTAGSTVRKSTVTKVTRKTKGKAVVR	240
DB	181	AGKAGARRKATVTKVHKIGVRRKTTARRTAGSTVRKSTVTKVTRKTKGKAVVR	240
QY	241	KSTVTRTARRPAGKTPGARRAGAKRR	269
DB	241	KSTVTRTARRPAGKTPGARRAGAKRR	269

RESULT 2
HC2X_CHLTR
ID HC2X_CHLTR STANDARD; PRT; 221 AA.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 14:39:24 ; Search time 40 Seconds

(without alignments)
647,058 Million cell updates/sec

Title: US-10-034-623-6

Sequence: 1 MGIEGGRGDMSENFVAFVCV.....RPAQKRTPGRAARAGAKRR 269

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1357	100.0	269	2	histone H1 DNA bin
2	220.5	16.2	223	2	histone H1 homolog
3	219	16.1	207	2	Hc2 nucleoprotein
4	217	16.0	207	2	histone H1 homolog
5	204	15.0	203	2	probable histone-1
6	203.5	15.0	172	2	histone-like prote
7	203.5	15.0	172	2	histone-like prote
8	201	14.8	328	2	salivary glue prot
9	199.5	14.7	182	2	histone H1 homolog
10	174	12.8	165	2	histone H1 homolog
11	169	12.5	218	2	histone H1 homolog
12	154	11.3	1390	2	histone-like DNA b
13	153	11.3	248	1	sperm tail-specific
14	152.5	11.2	288	1	histone H1, gonada
15	151.5	11.2	1052	1	histone H1 (clone
16	151.5	11.2	1701	2	kinetoplast-associ
17	148	10.9	135	2	probable erythrocy
18	148	10.9	229	2	histone H1 - Duplo
19	145.5	10.7	221	2	histone H1 - beta, e
20	142.5	10.5	282	2	histone H1 - com
21	142.5	10.5	284	2	histone H1 (clone
22	140.5	10.4	241	2	histone H1-II - Vo
23	139	10.2	1403	2	probable translati
24	138.5	10.2	384	2	outer membrane pro
25	138.5	10.2	1773	2	IGA-specific metal
26	138.5	10.2	1815	2	IGA-specific metal
27	136.5	10.1	421	2	histone H1 - Bsc
28	136	10.0	256	2	histone H1 - fruit
29	136	10.0	289	2	conserved hypothet

30	136	10.0	365	2	T35620	hypothetical prote
31	136	10.0	1203	2	S26650	DNA-binding protei
32	134	9.9	309	2	G83013	polyhydroxyalkanoe
33	132.5	9.8	214	2	G70673	probable hupb - My
34	132	9.7	206	2	S09388	histone H1 - sea u
35	132	9.7	952	2	T36664	hypothetical prote
36	131	9.7	485	2	T13479	probable DNA-bind
37	130.5	9.6	239	2	T34945	hypothetical lysin
38	129	9.5	239	2	T34733	hypothetical prote
39	129	9.5	289	2	A81779	probable DNA-bind
40	127	9.4	352	2	G82990	hypothetical prote
41	127	9.4	510	1	S41307	hypothetical lysin
42	127	9.4	528	2	B75310	arginate regulator
43	126.5	9.3	271	2	T06392	transcription init
44	126.5	9.3	1173	2	T31421	conserved hypothet
45	126	9.3	210	2	A25550	histone H1 - tom
						C-terminal domain-
						histone H1 - sea u

ALIGNMENTS

RESULT 1
T31309
histone H1 DNA binding protein - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T31309
R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of the
A:Reference number: Z20994; MUID:98422450; PMID:9748430
A:Accession: T31309
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-269 <SCH>
A:Cross-references: UNIPROT:O74057; EMBL:AF083072; NID:g3599393; PID:g3599395; PIDN:AF
A:Gene: hc2

Query Match 100.0%; Score 1357; DB 2; Length 269;
Best local similarity 100.0%; Pred. No. 1.7e-86;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGIEGGRGDMSENFVAFVCVACAGVTKDQEMKYVDGRVFRKECHARRGGQIRFENPVEQ	60
DB	1	MGIEGGRGDMSENFVAFVCVACAGVTKDQEMKYVDGRVFRKECHARRGGQIRFENPVEQ	60
QY	61	RVAEELKVDLIQMNQIAEMNRASGDGVHSSATSAAEAQRAELKVQIVQMNQIAEMN	120
DB	61	RVAEELKVDLIQMNQIAEMNRASGDGVHSSATSAAEAQRAELKVQIVQMNQIAEMN	120
QY	121	RAAPGKPAKKAAGTARPKSGKTVRRRTGRTGKAGARRKTVKRTARRKTTAKKA	180
DB	121	RAAPGKPAKKAAGTARPKSGKTVRRRTGRTGKAGARRKTVKRTARRKTTAKKA	180
QY	181	AGKAGARRKATVKTTHKKIGVRKTTARRTAGSTVRRKSTVTKTVARRKTKGKAVVR	240
DB	181	AGKAGARRKATVKTTHKKIGVRKTTARRTAGSTVRRKSTVTKTVARRKTKGKAVVR	240
QY	241	KSTVKTARRPAGRTPGRAARAGAKRR	269
DB	241	KSTVKTARRPAGRTPGRAARAGAKRR	269

RESULT 2
A36884
histone H1 homolog Hc2 - Chlamydia trachomatis
C:Species: Chlamydia trachomatis
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: A36884; JN0851
R:Brickman, T.J.; Barry III, C.E.; Hackstadt, T.
J. Bacteriol. 175, 4274-4281, 1993
A:Title: Molecular cloning and expression of hctB encoding a strain-variant chlamydial

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 18, 2004, 14:28:24 / Search time 132 Seconds

(without alignments)
658.885 Million cell updates/sec

Title: US-10-034-623-6

Perfect score: 1357

Sequence: 1 MHGIEGRGDMSENFVAFV.....RPAGKTPGARAPAGAKRR 269

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1360919 seqs, 323318874 residues

Total number of hits satisfying chosen parameters: 1360919

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA:
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1357	100.0	269	US-10-027-806-6	Sequence 6, Appl1
2	1357	100.0	269	US-10-034-623-6	Sequence 6, Appl1
3	1357	100.0	269	US-10-027-801-6	Sequence 6, Appl1
4	1357	100.0	269	US-10-029-120-6	Sequence 6, Appl1
5	208	15.3	376	US-10-156-761-9889	Sequence 9889, Ap
6	204	15.0	203	US-09-820-843A-7	Sequence 7, Appl1
7	203.5	15.0	172	US-09-820-843A-5	Sequence 5, Appl1
8	203.5	15.0	182	US-10-289-762-402	Sequence 402, App
9	177	13.0	217	US-10-156-761-10221	Sequence 10221, A
10	160.5	11.8	356	US-10-156-761-13658	Sequence 13658, A
11	159.5	11.8	329	US-10-282-122A-67699	Sequence 67699, A
12	155.5	11.5	212	US-10-282-122A-61735	Sequence 61735, A
13	154	11.3	969	US-10-282-122A-50758	Sequence 50758, A
14	152.5	11.2	301	US-10-437-963-182491	Sequence 182491, A

15	149.5	11.0	289	US-10-437-963-144785	Sequence 144785, A
16	148	10.9	223	US-10-051-643-201	Sequence 201, App
17	148	10.9	223	US-10-205-979-52	Sequence 52, Appl1
18	148	10.9	1387	US-10-156-761-13000	Sequence 13000, A
19	147.5	10.9	260	US-10-437-963-107671	Sequence 107671, A
20	146.5	10.8	366	US-10-156-761-104871	Sequence 104871, A
21	145.5	10.7	184	US-10-424-599-188977	Sequence 188977, A
22	140	10.3	337	US-10-437-963-157680	Sequence 157680, A
23	139	10.2	19608	US-10-084-846A-8	Sequence 8, Appl1
24	138.5	10.2	428	US-10-282-122A-55748	Sequence 55748, A
25	138.5	10.2	446	US-10-437-963-127362	Sequence 127362, A
26	138.5	10.2	1773	US-10-282-122A-65817	Sequence 65817, A
27	138	10.2	279	US-10-181-071-7	Sequence 7, Appl1
28	137.5	10.1	942	US-10-156-761-12155	Sequence 12155, A
29	137	10.1	2263	US-10-408-765A-2231	Sequence 2231, Ap
30	136.5	10.1	421	US-10-282-122A-56483	Sequence 56483, A
31	136	10.0	211	US-10-437-963-174659	Sequence 174659, A
32	136	10.0	228	US-10-156-761-9425	Sequence 9425, Ap
33	136	10.0	321	US-09-864-761-33718	Sequence 33718, A
34	136	10.0	361	US-09-864-761-33614	Sequence 33614, A
35	136	10.0	361	US-09-864-761-33614	Sequence 33614, A
36	136	10.0	1047	US-10-029-386-33522	Sequence 33522, A
37	136	10.0	1087	US-10-029-386-32407	Sequence 32407, A
38	135.5	10.0	202	US-10-437-963-151115	Sequence 151115, A
39	135.5	10.0	204	US-10-437-963-151115	Sequence 151115, A
40	135.5	10.0	838	US-10-156-761-10342	Sequence 10342, A
41	135	9.9	221	US-10-424-599-252204	Sequence 252204, A
42	134.5	9.9	315	US-10-437-963-200938	Sequence 200938, A
43	134.5	9.9	503	US-10-282-122A-50517	Sequence 50517, A
44	134	9.9	251	US-10-437-963-180341	Sequence 180341, A
45	134	9.9	309	US-09-820-843A-24	Sequence 24, Appl1

ALIGNMENTS

RESULT 1	US-10-027-806-6	Sequence 6, Application US/10027806
Publication No.	US20020160476A1	
GENERAL INFORMATION:		
APPLICANT:	Swanson, Ronald V.	
APPLICANT:	Feldman, Robert A.	
APPLICANT:	Schleper, Christa	
TITLE OF INVENTION:	NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM	
FILE REFERENCE:	DCORP.002A	
CURRENT APPLICATION NUMBER:	US/10/027,806	
CURRENT FILING DATE:	2001-12-21	
PRIOR APPLICATION NUMBER:	EARLIER APPLICATION NUMBER: 09/408,020	
PRIOR FILING DATE:	EARLIER FILING DATE: 1999-09-23	
NUMBER OF SEQ ID NOS:	123	
SOFTWARE:	FastSeq for Windows Version 3.0	
SEQ ID NO 6		
LENGTH:	269	
TYPE:	PRT	
ORGANISM:	Cenarchaeum symbiosum	
US-10-027-806-6		
Query Match	100.0%; Score 1357; DB 13; Length 269;	
Best Local Similarity	100.0%; Pred. No. 3.5e-113;	
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MHGIEGRGDMSENFVAFVCAACAGCTTCKEMKVDVRFHKECHAHGQIRPNPEVEQ 60	
DB	1 MHGIEGRGDMSENFVAFVCAACAGCTTCKEMKVDVRFHKECHAHGQIRPNPEVEQ 60	
QY	61 RVAELVYDILQENQLAENNRASGDGVHSATSAAAEQRAELKVLQVQENQLAENN 120	
DB	61 RVAELVYDILQENQLAENNRASGDGVHSATSAAAEQRAELKVLQVQENQLAENN 120	
QY	121 RPAKGPAPAKKAAGKTARRKSGKTVRRKTKRTAKKAAGARRKTVKRTARRKTTAKKA 180	
DB	121 RPAKGPAPAKKAAGKTARRKSGKTVRRKTKRTAKKAAGARRKTVKRTARRKTTAKKA 180	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 18, 2004, 14:19:22 ; Search time 40 Seconds
(without alignments)
445.989 Million cell updates/sec

Title: US-10-034-623-6

Perfect score: 1357

Sequence: 1 MHGIEGGGDMSENFVAFCV.....RPAKRTPEGRAPRAAKRR 269

Scoring table:

BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCtUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfltest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1357	100.0	269	4	US-09-408-020-6
2	203.5	15.0	182	4	US-09-198-452A-402
3	148	10.9	223	3	US-09-095-855-201
4	148	10.9	223	4	US-09-205-426-201
5	147	10.8	246	4	US-09-252-991A-16586
6	137.5	10.1	492	4	US-09-252-991A-28339
7	135.5	10.0	315	4	US-09-252-991A-20553
8	132.5	9.8	214	3	US-09-041-889-27
9	132.5	9.8	214	4	US-09-417-264-27
10	132.5	9.8	443	4	US-09-252-991A-26132
11	132.5	9.8	722	4	US-09-252-991A-24102
12	132	9.7	434	1	US-08-097-830E-3
13	132	9.7	434	4	US-08-456-112B-3
14	132	9.7	1523	4	US-09-538-092-955
15	130.5	9.6	491	4	US-09-489-039A-7836
16	130.5	9.6	726	4	US-09-392-714-21
17	130	9.6	536	4	US-09-252-991A-16754
18	129.5	9.5	212	4	US-09-489-039A-15074
19	129.5	9.5	437	4	US-09-252-991A-25942
20	129.5	9.5	469	4	US-09-489-039A-13565
21	129	9.5	283	4	US-09-270-767-55111
22	129	9.5	283	4	US-09-270-767-55111
23	128.5	9.5	543	4	US-09-252-991A-31103
24	128	9.4	195	4	US-09-252-991A-25427
25	127.5	9.4	561	4	US-09-252-991A-23080
26	127.5	9.4	757	4	US-09-252-991A-23569
27	127.5	9.4	1213	3	US-09-413-814-79

28	127	9.4	399	4	US-09-252-991A-22853	Sequence 22853, A
29	126.5	9.3	240	4	US-09-252-991A-16592	Sequence 16592, A
30	126.5	9.3	310	4	US-09-252-991A-28514	Sequence 28514, A
31	126	9.3	294	4	US-09-270-767-36084	Sequence 36084, A
32	126	9.3	294	4	US-09-270-767-51301	Sequence 51301, A
33	126	9.3	728	4	US-09-252-991A-23613	Sequence 23613, A
34	125	9.2	316	4	US-09-252-991A-32957	Sequence 32957, A
35	125	9.2	488	4	US-09-252-991A-30531	Sequence 30531, A
36	125	9.2	755	4	US-09-252-991A-28386	Sequence 28386, A
37	124.5	9.2	458	4	US-09-252-991A-32991	Sequence 32991, A
38	124	9.1	329	4	US-09-252-991A-26389	Sequence 26389, A
39	124	9.1	614	4	US-09-252-991A-20210	Sequence 20210, A
40	124	9.1	631	4	US-09-252-991A-30909	Sequence 30909, A
41	124	9.1	1041	4	US-09-252-991A-25637	Sequence 25637, A
42	123	9.1	235	4	US-09-270-767-36681	Sequence 36681, A
43	123	9.1	235	4	US-09-270-767-51898	Sequence 51898, A
44	123	9.1	361	4	US-09-252-991A-28125	Sequence 28125, A
45	122.5	9.0	299	4	US-09-252-991A-21789	Sequence 21789, A

ALIGNMENTS

```

RESULT 1
US-09-408-020-6
Sequence 6, Application US/09408020
Patent No. 6632937
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
APPLICANT: Schlieper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAUM SYMBIOSUM
FILE REFERENCE: DCOB.002A
CURRENT APPLICATION NUMBER: US/09/408,020
CURRENT FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: 60/102,294
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 269
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-09-408-020-6

Query Match      100.0%; Score 1357; DB 4; Length 269;
Best Local Similarity 100.0%; Pred No. 2.6e-137;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MHGIEGGGDMSENFVAFCVACARGVTDGKRVFHKCHARGGQIRFPNPEVQ 60
DB      1 MHGIEGGGDMSENFVAFCVACARGVTDGKRVFHKCHARGGQIRFPNPEVQ 60

QY      61 RFAELKVDIQRNOIAENRASSGGVSSSTSAEAFQHAELKVQIVQRNOIAEN 120
DB      61 RFAELKVDIQRNOIAENRASSGGVSSSTSAEAFQHAELKVQIVQRNOIAEN 120

QY      121 RPAKPPARKKAAGTAAKRSKKTGRTGRTGKAGARAKTTVKTARTKTTAKA 180
DB      121 RPAKPPARKKAAGTAAKRSKKTGRTGRTGKAGARAKTTVKTARTKTTAKA 180

QY      181 AGRKAGARKKATVTKTKTGVRKRTTARTAGSKTVRKSTVTKRTGKAVVR 240
DB      181 AGRKAGARKKATVTKTKTGVRKRTTARTAGSKTVRKSTVTKRTGKAVVR 240

QY      241 KSTVKTARPPAKRTPEGRAPRAAKRR 269
DB      241 KSTVKTARPPAKRTPEGRAPRAAKRR 269

RESULT 2
US-09-198-452A-402
Sequence 402, Application US/09198452A

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 18, 2004, 14:30:38 ; Search time 160 Seconds
(without alignments)
603.114 Million cell updates/sec

Title: US-10-034-623-6

Perfect score: 1357

Sequence: 1 MHGIEGGRGDMSENFVAFCV.....RPAGRKTPGARAARAGAKRR 269

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseq1980s:*\n2: Geneseq1980s:*\n3: Geneseq2000s:*\n4: Geneseq2000s:*\n5: Geneseq2002s:*\n6: Geneseq2003as:*\n7: Geneseq2003bs:*\n8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1357	100.0	269	AAV90914	AAV90914 Cenarchae
2	226	16.7	295	AAV37179	AAV37179 Protein i
3	204	15.0	203	ABO23506	ABO23506 Chlamydia
4	203.5	15.0	172	ABO23504	ABO23504 Chlamydia
5	182	15.0	182	AAV34984	AAV34984 Amino act
6	159.5	11.8	329	ABU33775	ABU33775 Protein e
7	155.5	11.3	212	ABU33811	ABU33811 Protein e
8	154	11.3	969	ABU22834	ABU22834 Protein e
9	153	11.3	290	AAV38693	AAV38693 Neisseria
10	151.5	11.2	540	ADML19717	ADML19717 Protein e
11	148	10.9	223	AAV14928	AAV14928 Amino act
12	148	10.9	223	ABP70903	ABP70903 Mycobacte
13	147	10.8	246	ABO67840	ABO67840 Pseudomon
14	146	10.8	252	AAV74897	AAV74897 Neisseria
15	145.5	10.7	583	ABG09005	ABG09005 Novel hum
16	144.5	10.6	267	AAV38692	AAV38692 Neisseria
17	144.5	10.6	267	ABG20340	ABG20340 Novel hum
18	141	10.4	1275	ABG20340	ABG20340 Novel hum
19	140	10.3	289	ABP80636	ABP80636 N. gonorr
20	139	10.2	467	ABP69558	ABP69558 Human pol
21	139	10.2	1098	ABG20365	ABG20365 Novel hum
22	138.5	10.2	428	ABU27824	ABU27824 Protein e
23	138.5	10.2	1773	ABU37893	ABU37893 Protein e
24	138	10.2	279	AAU03592	AAU03592 Human DNA
25	138	10.2	467	ABB94309	ABB94309 Human pro

26	137.5	10.1	392	2	AAV96808	AAV96808 Neisseria
27	137.5	10.1	492	7	ABO79593	ABO79593 Pseudomon
28	137	10.1	266	3	AAV74898	AAV74898 Neisseria
29	137	10.1	956	6	ABO53053	ABO53053 Human put
30	137	10.1	2263	7	ADJ70425	ADJ70425 Human hea
31	136.5	10.1	421	6	ABU28559	ABU28559 Protein e
32	136	10.0	289	2	AAV38690	AAV38690 Neisseria
33	136	10.0	321	4	AAV14001	AAV14001 Peptide #
34	136	10.0	321	4	ABV32946	ABV32946 Peptide #
35	136	10.0	321	4	AAV26407	AAV26407 Peptide #
36	136	10.0	321	4	ABV27775	ABV27775 Human pep
37	136	10.0	321	4	ABV18420	ABV18420 Protein #
38	136	10.0	321	4	AAV6130	AAV6130 Human bon
39	136	10.0	321	4	AAV57747	AAV57747 Human bra
40	136	10.0	321	4	AAV47798	AAV47798 Human liv
41	136	10.0	321	4	AAV01742	AAV01742 Peptide #
42	136	10.0	321	5	ABG35780	ABG35780 Human rep
43	136	10.0	361	4	AAV14436	AAV14436 Peptide #
44	136	10.0	361	4	AAV13889	AAV13889 Peptide #
45	136	10.0	361	4	ABV32834	ABV32834 Peptide #

ALIGNMENTS

RESULT 1

AAV90914 standard; protein; 269 AA.

AAV90914;

30-AUG-2000 (first entry)

Cenarchaeum symbiosum open reading frame protein sequence SEQ ID NO:6.

Cenarchaeum symbiosum; non-thermophilic; crenarchaeote; physiology; characterisation; archae; therapeutic; industrial; laboratory.

Cenarchaeum symbiosum.

WO200018909-A2.

06-APR-2000.

29-SEP-1999; 99WO-US022752.

29-SEP-1998; 98US-0102294P.

(DIVE-) DIVERSA CORP.

Swanson RV, Feldman RA, Schleper C;

WPI; 2000-293148/25.

N-PSDB; AA55189.

New nucleic acids and proteins isolated from the non-thermophilic crenarchaeote Cenarchaeum symbiosum, useful in characterizing the physiology of these archae and in therapeutic, industrial or laboratory techniques.

Claim 24; Page 111-112; 210pp; English.

AA55186 to AA55226 and AAV90913 to AAV90951 represent nucleic acids and proteins isolated from the non-thermophilic crenarchaeote Cenarchaeum symbiosum. The nucleic acids and proteins identified in the present invention are useful in characterizing the physiology of these archae and can be used in therapeutic, industrial or laboratory techniques. AA55227 to AA55260 represent promoter sequences from Cenarchaeum symbiosum. AA55261 to AA55269 represent PCR primers and probes used in examples from the present invention

Sequence 269 AA;

SO